

Figure 4

A

FIX 176834-GTCTGCAAC**ATG**CGGCTGTGTGCGGTGTGGCTGTCTGTTTGTCTGTGCGCCGTGGTGCTG
RACE_95_3 GTCTGCAAC**ATG**CGGCTGTGTGCGGTGTGGCTGTCTGTTTGTCTGTGCGCCGTGGTGCTG
RACE_95_8 GTCTGCAAC**ATG**CGGCTGTGTGCGGTGTGGCTGTCTGTTTGTCTGTGCGCCGTGGTGCTG
RACE_95_11 GTCTGCAAC**ATG**CGGCTGTGTGCGGTGTGGCTGTCTGTTTGTCTGTGCGCCGTGGTGCTG

FIX GTTCAGTGCCAGCGGGAGACCGCAGAAAAA--CGATTATTACCGAGTACCGCATTACT
RACE_95_3 GTTCAGTGCCAGCGGGAGACCGCAGAAAAA--CGATTATTACCGAGTACCGCATTACT
RACE_95_8 GTTCAGTGCCAGCGGGAGACCGCAGAAAAA--CGATTATTACCGAGTACCGCATTACT
RACE_95_11 GTTCAGTGCCAGCGGGAGACCGCAGAAAAAACGATTATTACCGAGTACCGCATTACT

FIX GGGACGCGTGCTCTCGCGCGCTGCCTGACCAAACCCGTTACAAGTATGTGGAACAGCTCG
RACE_95_3 GGGACGCGTGCTCTCGCGCGCTGCCTGACCAAACCCGTTACAAGTATGTGGAACAGCTCG
RACE_95_8 GGGACGCGTGCTCTCGCGCGCTGCCTGACCAAACCCGTTACAAGTATGTGGAACAGCTCG
RACE_95_11 GGGACGCGTGCTCTCGCGCGCTGCCTGACCAAACCCGTTACAAGTATGTGGAACAGCTCG

FIX TGGACCTCACGTTGAACTACCACTACGATGCGAGCCACGGCTTGGACAACCTTTGACGTGC
RACE_95_3 TGGACCTCACGTTGAACTACCACTACGATGCGAGCCACGGCTTGGACAACCTTTGACGTGC
RACE_95_8 TGGACCTCACGTTGAACTACCACTACGATGCGAGCCACGGCTTGGACAACCTTTGACGTGC
RACE_95_11 TGGACCTCACGTTGAACTACCACTACGATGCGAGCCACGGCTTGGACAACCTTTGACGTGC

FIX TCAAGAG**GTG**AGGGTACGCGCTAAAGGTGTATGACAACGGGAAGGTAAGGGCGAACGGGT
RACE_95_3 TCAAGAG-----
RACE_95_8 TCAAGAG**GTG**AGGGTACGCGCTAAAGGTGTATGACAACGGGAAGGTAAGGGCGAACGGGT
RACE_95_11 TCAAGAG-----

FIX AACGGGTAGGTAACCGCATGGGGTGTGAAATGACGTTCCGAACCTGTGCTTGC**GA**AATCA
RACE_95_3 -----AATCA
RACE_95_8 AACGGGTAGGTAACCGCATGGGGTGTGAAATGACGTTCCGAACCTGTGCTTGC**GA**AATCA
RACE_95_11 -----AATCA

FIX ACGTGACCGAGGTGTGCTGCTCATCAGCGACTTTAGACGTCAGAACCGTCGCGGCGGCA
RACE_95_3 ACGTGACCGAGGTGTGCTGCTCATCAGCGACTTTATACGTCAGAACCGTCGCGGCGGCA
RACE_95_8 ACGTGACCGAGGTGTGCTGCTCATCAGCGACTTTAGACGTCAGAACCGTCGCGGCGGCA
RACE_95_11 ACGTGACCGAGGTGTGCTGCTCATCAGCGACTTTAGACGTCAGAACCGTCGCGGCGGCA

FIX CCAACAAAAGGACCACGTTCAACGCCGCCGGTTTCGCTGGCGCCTCACGCCCGGAGCCTCG
RACE_95_3 CCAACAAAAGGACCACGTTCAACGCCGCCGGTTTCGCTGGCGCCTCACGCCCGGAGCCTCG
RACE_95_8 CCAACAAAAGGACCACGTTCAACGCCGCCGGTTTCGCTGGCGCCTCACGCCCGGAGCCTCG
RACE_95_11 CCAACAAAAGGACCACGTTCAACGCCGCCGGTTTCGCTGGCGCCTCACGCCCGGAGCCTCG

FIX AGTTCAGCGTGCGGCTCTTTGCCAACT**AG**CCTGCGTCA-176346
RACE_95_3 AGTTCAGCGTGCGGCTCTTTGCCAACT**AG**CCTGCGTCA
RACE_95_8 AGTTCAGCGTGCGGCTCTTTGCCAACT**AG**CCTGCGTCA
RACE_95_11 AGTTCAGCGTGCGGCTCTTTGCCAACT**AG**CCTGCGTCA

B

FIX 175631-CCGCGCGTCA**ATG**AGTCCCAAAAACCTGACGCCGTTCTTGACGGCGTTGTGGCTGCTATTG
RACE_95_3 CCGTGCGTCA**ATG**AGTCCCAAAAACCTGACGCCGTTCTTGACGGCGTTGTGGCTGCTATTG
RACE_95_8 CCGCGCGTCA**ATG**AGTCCCAAAAACCTGACGCCGTTCTTGACGGCGTTGTGGCTGCTATTG
RACE_95_13 CCGCGCGTCA**ATG**AGTCCCAAAAACCTGACGCCGTTCTTGACGGCGTTGTGGCTGCTATTG

FIX7 GGTACAGCCCGGTGCGCGGGTACGCGCAGAAGAATGTTGCGAATTCATAAACGTCAAC
RACE_95_3 GGTACAGCCCGGTGCGCGGGTACGCGCAGAAGAATGTTGCGAATTCATAAACGTCAAC
RACE_95_8 GGTACAGCCCGGTGCGCGGGTACGCGCAGAAGAATGTTGCGAATTCATAAACGTCAAC
RACE_95_11 GGTACAGCCCGGTGCGCGGGTACGCGCAGAAGAATGTTGCGAATTCATAAACGTCAAC

FIX CACCCGCCGGAACGCTGTTACGATTTCAAATGTGCAATCGCTTCACCGTCGCG**GA**ACGTA
RACE_95_3 CACCCGCCGGAACGCTGTTACGATTTCAAATGTGCAATCGCTTCACCGTCGCG**GA**ACGTA
RACE_95_8 CACCCGCCGGAACGCTGTTACGATTTCAAATGTGCAATCGCTTCACCGTCGCG**GA**ACGTA
RACE_95_11 CACCCGCCGGAACGCTGTTACGATTTCAAATGTGCAATCGCTTCACCGTCGCG-----

Fig. 4 UL131-128 mRNA processing – Panels (A-C) compare FIX-BAC DNA sequence (numbered according to Chee et al.) to a set of cDNA sequences from RACE clones 95-3, 95-8 and 95-11 (A) UL131 region, (B) UL128 region, (C) UL131-128 transcripts 3' end. Start codons, stop codons and the polyA site are in bold face, mRNA processing signals (splice donor sequence, splice acceptor sequence, AATAAA signal) are grey-shaded.

Figure 7

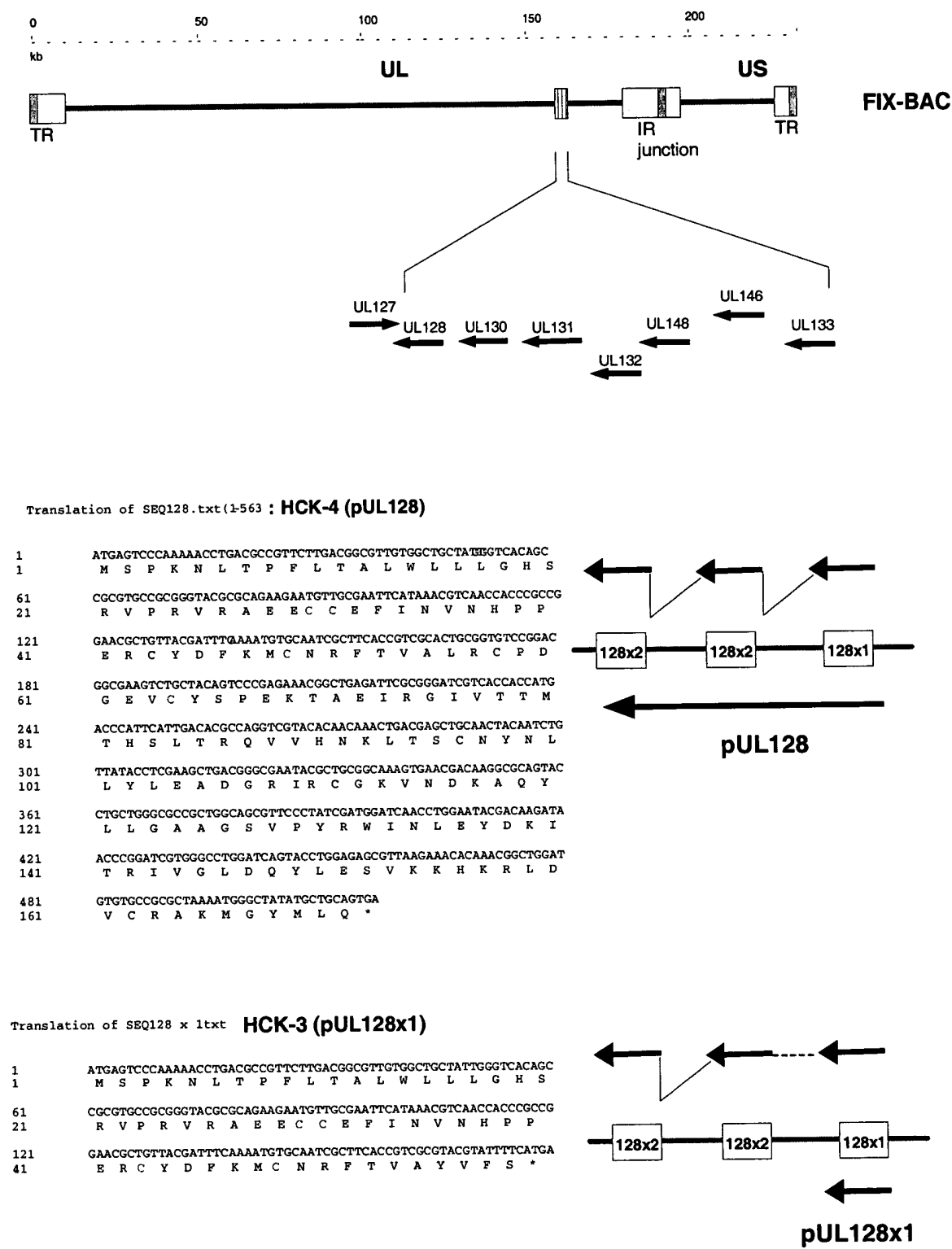


Fig. 7 Scheme of the differentially spliced transcripts of the UL131-128 region. Upper panel SEQUAL128B and predicted open reading frame (orf) pUL128 (HCK-4). Lower panel SEQUAL128A and predicted orf UL128x1 (HCK-3).

Figure 8

Northern Blot Analysis

RVFIX, RVFIX mutants and laboratory strains:

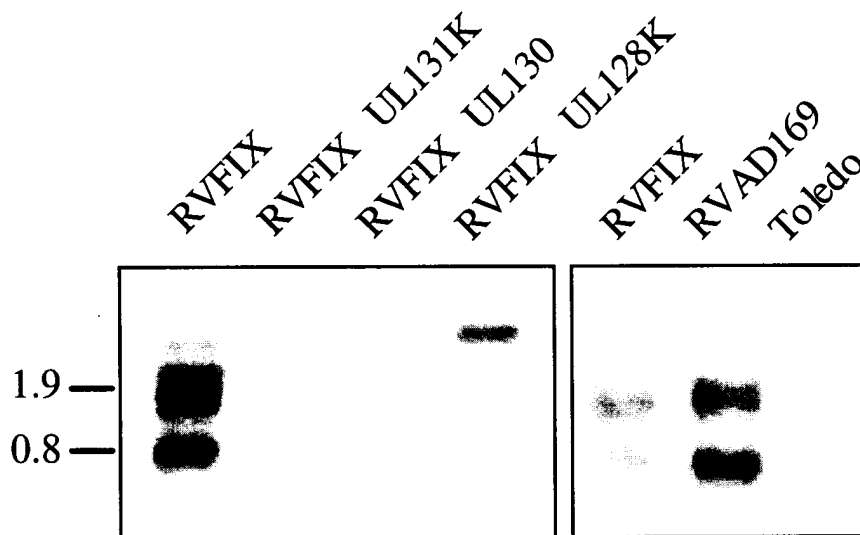


Fig. 8 mRNA was prepared from RVFIX-infected fibroblasts 4 days p.i. using Rneasy Mini, QIAshredder and Oligotex mRNA Mini kits according to the manufacturer's guidance (Qiagen). For Northern blotting, 1 μ g RNA was electrophoresed on an agarose gel according to the MOPS-formaldehyde protocol and blotted onto Hybond N+ membranes (Amersham Pharmacia). Blots were hybridized with a UL131-128 specific probe.

Figure 9

Comparison RACE clone 95-3 – FIX genomic sequence

Upper line: SEQFIX UL131-128.txt, from 10 to 1977

Lower line: SEQKlon95-3.txt, from 1 to 1741

SEQFIX UL131-128.txt:SEQKlon95-3.txt identity= 99.66%(1735/1741)
gap=11.94%(236/1977)

```
1      GTCTGCAACATGCGGCTGTGTCTGGGTGTGGCTGTCTGTTTGTCTGTGCGCCGTGGTGTCTG
      |||||||
1      .....ATGCGGCTGTCTCGGGTGTGGCTGTCTGTTTGTCTGTGCGCCGTGGTGTCTG

61     GGTCAGTGCCAGCGGGAGACCGCAGAAAAAACGATTATTACCGAGTACCGCATTACTGG
      |||||||
52     GGTCAGTGCCAGCGGGAGACCGCAGAAAAAACGATTATTACCGAGTACCGCATTACTGG

121    GACGCGTGCTCTCGCGCGCTGCCTGACCAAACCCGTTACAAGTATGTGGAACAGCTCGTG
      |||||||
112    GACGCGTGCTCTCGCGCGCTGCCTGACCAAACCCGTTACAAGTATGTGGAACAGCTCGTG

181    GACCTCACGTTGAACTACCACTACGATGCGAGCCACGGCTTGGACAACCTTTGACGTGCTC
      |||||||
172    GACCTCACGTTGAACTACCACTACGATGCGAGCCACGGCTTGGACAACCTTTGACGTGCTC

241    AAGAGGTGAGGGTACGCGCTAAAGGTGTATGACAACGGGAAGGTAAGGGCGAACGGGTAA
      |||||
232    AAGAG.....

301    CGGGTAGGTAACCGCATGGGGTGTGAAATGACGTTTCGGAACCTGTGCTTGCAGAATCAAC
      |||||||
235    .....AATCAAC

361    GTGACCGAGGTGTCGTTGCTCATCAGCGACTTTAGACGTCAGAACCGTCGCGGCGGCACC
      |||||||
244    GTGACCGAGGTGTCGTTGCTCATCAGCGACTTTATACGTCAGAACCGTCGCGGCGGCACC

421    AACAAAAGGACCACGTTCAACGCCGCCGTTTCGCTGGCGCCTCACGCCCGGAGCCTCGAG
      |||||||
304    AACAAAAGGACCACGTTCAACGCCGCCGTTTCGCTGGCGCCTCACGCCCGGAGCCTCGAG

481    TTCAGCGTGCGGCTCTTTGCCAACTAGCCTGCGTCACGGGAAATAATATGCTACGGCTTC
      |||||||
364    TTCAGCGTGCGGCTCTTTGCCAACTAGCCTGCGTCACGGGAAATAATATGCTACGGCTTC

541    TGCTTCGTCACCACTTTCACTGCCTGCTTCTGTGCGCGGTTTGGGCAACGCCCTGTCTGG
      |||||||
424    TGCTTCGTCACCACTTTCACTGCCTGCTTCTGTGCGCGGTTTGGGCAACGCCCTGTCTGG

601    CGTCTCCGTGGTTTCACGCTAACGGCGAACCAGAATCCGTCCCCGCCATGGTCTAAACTGA
      |||||||
484    CGTCTCCGTGGTTTCACGCTAACGGCGAACCAGAATCCGTCCCCGCCATGGTCTAAACTGA

661    CGTATCCCAAACCGCATGACGCGGCGACGTTTTACTGTCCTTTTCTCTATCCCTCGCCCC
      |||||||
544    CGTATCCCAAACCGCATGACGCGGCGACGTTTTACTGTCCTTTTCTCTATCCCTCGCCCC
```

721 CACGGTCCCCCTCGCAATTCCCGGGGTTCAGCGGGTATCAACGGGTCCCGAGTGTGCGA
|||
604 CACGGTCCCCCTCGCAATTCCCGGGGTTCAGCGGGTATCAACGGGTCCCGAGTGTGCGA

781 ACGAGACCCTGTATCTGCTGTACAACCGGGAAGGCCAGACCTTGGTGGAGAGAAGCTCCA
|||
664 ACGAGACCCTGTATCTGCTGTACAACCGGGAAGGCCAGACCTTGGTGGAGAGAAGCTCCA

841 CCTGGGTGAAAAAGGTGATCTGGTATCTGAGCGGTCGCAATCAGACCATCCTCCAACGGA
|||
724 CCTGGGTGAAAAAGGTGATCTGGTATCTGAGCGGTCGCAATCAGACCATCCTCCAACGGA

901 TGCCCCGAACGGCTTCGAAACCGAGCGACGGAAACGTGCAGATCAGCGTGGAAAGACGCCA
|||
784 TGCCCCGAACGGCTTCGAAACCGAGCGACGGAAACGTGCAGATCAGCGTGGAAAGACGCCA

961 AGATTTTGGAGCGCACATGGTGCCCAAGCAGACCAAGCTGCTACGTTTCGTCGTCACCG
|||
844 AGATTTTGGAGCGCACATGGTGCCCAAGCAGACCAAGCTGCTACGTTTCGTCGTCACCG

1021 ATGGCACACGTTATCAGATGTGTGTGATGAAACTGGAGAGCTGGGCCCACGTCTTCCGGG
|||
904 ATGGCACACGTTATCAGATGTGTGTGATGAAACTGGAGAGCTGGGCCCACGTCTTCCGGG

1081 ACTACAGCGTGTCTTTTCAGGTGCGATTGACGTTACCGAGGCCAATAACCAGACTTACA
|||
964 ACTACAGCGTGTCTTTTCAGGTGCGATTGACGTTACCGAGGCCAATAACCAGACTTACA

1141 CTTTCTGCACCCATCCCAATCTCATCGTTTGAGCCCGTCGCGCGCGCAGGGAATTTTGAA
|||
1024 CTTTCTGCACCCATCCCAATCTCATCGTTTGAGCCCGTCGCGCGCGCAGGGAATTTTGAA

1201 AACC CGCGTCATGAGTCCCAAAAACCTGACGCCGTTCTTGACGGCGTTGTGGCTGCTAT
|||
1084 AACC CGCGTCATGAGTCCCAAAAACCTGACGCCGTTCTTGACGGCGTTGTGGCTGCTAT

1261 TGGGTACAGCCGCGTGCCGCGGGTACGCGCAGAAGAATGTTGCGAATTCATAAACGTC A
|||
1144 TGGGTACAGCCGCGTGCCGCGGGTACGCGCAGAAGAATGTTGCGAATTCATAAACGTC A

1321 ACCACCCGCCGGAACGCTGTTACGATTTCAAATGTGCAATCGCTTCACCGTCGCGTACG
|||
1204 ACCACCCGCCGGAACGCTGTTACGATTTCAAATGTGCAATCGCTTCACCGTCGCGTACG

1381 TATTTTCATGATTGTCTGCGTTCTGTGGTGCGTCTGGATTTGTCTCTCGACGTTTCTGAT
|||
1264 TATTTTCATGATTGTCTGCGTTCTGTGGTGCGTCTGGATCTGTCTCTCGACGTTTCTGAT

1441 AGCCATGTTCCATCGACGATCCTCGGGAATGCCAGAGTAGATTTTCATGAATCCACAGGC
|||
1324 AGCCATGTTCCATCGACGATCCTCGGGAATGCCAGAGTAGATTTTCATGAATCCACAGGC

1501 TGCGGTGTCCGGACGGCGAAGTCTGCTACAGTCCCGAGAAAACGGCTGAGATTGCGGGGA
|||
1384 TGCGGTGTCCGGACGGCGAAGTCTGCTACAGTCCCGAGAAAACGGCTGAGATTGCGGGGA

1561 TCGTCACCACCATGACCCATTCAATTGACACGCCAGGTCGTACACAACAAACTGACGAGCT
|||
1444 TCGTCACCACCATGACCCATTCAATTGACACGCCAGGTCGTACACAACAAACTGACGAACT

1621 GCAACTACAATCCGTAAGTCTCTTCTCGAGGGCCTTACAGCCTATGGGAAAGTAAGACA
|||
1504 GCAACTACAATCC.....

1681 GAGGGACAAAACATCATTAATAAAAAAGTCTAATTTACGTTTGTACCCCCCTTCCCC
1517
1741 TCCGTGTTGTAGGTTATACCTCGAAGCTGACGGGCGAATACGCTGCGGCAAAGTGAACGA
|||
1517GTTATACCTCGAAGCTGACGGGCGAATACGCTGCGGCAAAGTGAACGA
1801 CAAGGCGCAGTACCTGCTGGGCGCCGCTGGCAGCGTTCCCTATCGATGGATCAACCTGGA
|||
1565 CAAGGCGCAGTACCTGCTGGGCGCCGCTGGCAGCGTTCCCTATCGATGGATCAACCTGGA
|||
1861 ATACGACAAGATAACCCGGATCGTGGGCCCTGGATCAGTACCTGGAGAGCGTTAAGAAACA
|||
1625 ATACGACAAGATAACCCGGATCGTGGGCCCTGGATCAGTACCTGGAGAGCGTTAAGAAACA
|||
1921 CAAACGGCTGGATGTGTGCCGCGCTAAAATGGGCTATATGCTGCAGTGAATAATAAA
|||
1685 CAAACGGCTGGATGTGTGCCGCGCTAAAATGGGCTATATGCTGCAGTGAATAATAAA
|||

Translation of SEQKlon95-3.txt: HCK-1 (pUL131)

1 ATGCGGCTGTCTCGGGTGTGGCTGTCTGTTGTCTGTGCGCCGTGGTGCTGGGTCACTGC
1 M R L S R V W L S V C L C A V V L G Q C
61 CAGCGGGAGACCGCAGAAAAAACGATTATTACCGAGTACCGCATTACTGGGACGCGTGC
21 Q R E T A E K N D Y Y R V P H Y W D A C
121 TCTCGCGCGCTGCCTGACCAAACCCGTTACAAGTATGTGGAACAGCTCGTGGACCTCACG
41 S R A L P D Q T R Y K Y V E Q L V D L T
181 TTGAACTACCACTACGATGCGAGCCACGGCTTGGACAACCTTGACGTGCTCAAGAGAATC
61 L N Y H Y D A S H G L D N F D V L K R I
241 AACGTGACCGAGGTGTCGTTGCTCATCAGCGACTTTATACGTCAGAACCGTCGCGGCGGC
81 N V T E V S L L I S D F I R Q N R R G G
301 ACCAACAAAAGGACCACGTTCAACGCCGCCGTTTCGCTGGCGCCTCACGCGCGGAGCCTC
101 T N K R T T F N A A G S L A P H A R S L
361 GAGTTCAGCGTGGGCTCTTTGCCAAGTAG
121 E F S V R L F A N *

Figure 10

Comparison RACE clone 95-8 –FIX genomic sequence

Upper line: SEQFIX UL131-128.txt, from 10 to 1977

Lower line: SEQKlon95-8.txt, from 1 to 1849

SEQFIX UL131-128.txt:SEQKlon95-8.txt identity= 99.78%(1845/1849)
gap=6.47%(128/1977)

```
1      GTCTGCAACATGCGGCTGTGTCTGGGTGTGGCTGTCTGTTGTCTGTGCGCCGTGGTGCTG
      |||
1      .....ATGCGGCTGTGTCTGGGTGTGGCTGTCTGTTGTCTGTGCGCCGTGGTGCTG

61     GGTCAGTGCCAGCGGGAGACCGCAGAAAAAACGATTATTACCGAGTACCGCATTACTGG
      |||
52     GGTCAGTGCCAGCGGGAGACCGCAGAAAAAACGATTATTACCGAGTACCGCATTACTGG

121    GACGCGTGCTCTCGCGCGCTGCCTGACCAAACCCGTTACAAGTATGTGGAACAGCTCGTG
      |||
112    GACGCGTGCTCTCGCGCGCTGCCTGACCAAACCCGTTACAAGTATGTGGAACAGCTCGTG

181    GACCTCACGTTGAACTACCACTACGATGCGAGCCACGGCTTGGACAACCTTTGACGTGCTC
      |||
172    GACCTCACGTTGAACTACCACTACGATGCGAGCCACGGCTTGGACAACCTTTGACGTGCTC

241    AAGAGGTGAGGGTACGCGCTAAAGGTGTATGACAACGGGAAGGTAAGGGCGAACGGGTAA
      |||
232    AAGAGGTGAGGGTACGCGCTAAAGGTGTATGACAACGGGAAGGTAAGGGCGAACGGGTAA

301    CGGGTAGGTAACCGCATGGGGTGTGAAATGACGTTTCGGAACCTGTGCTTGCAGAATCAAC
      |||
292    CGGGTAGGTAACCGCATGGGGTGTGAAATGACGTTTCGGAACCTGTGCTTGCAGAATCAAC

361    GTGACCGAGGTGTCGTTGCTCATCAGCGACTTTAGACGTCAGAACCGTCGCGGCGGCACC
      |||
352    GTGACCGAGGTGTCGTTGCTCATCAGCGACTTTAGACGTCAGAACCGTCGCGGCGGCACC

421    AACAAAAGGACCACGTTCAACGCCGCCGGTTCGCTGGCGCCTCACGCCCGGAGCCTCGAG
      |||
412    AACAAAAGGACCACGTTCAACGCCGCCGGTTCGCTGGCGCCTCACGCCCGGAGCCTCGAG

481    TTCAGCGTGCGGCTCTTTGCCAACTAGCCTGCGTCACGGGAAATAATATGCTACGGCTTC
      |||
472    TTCAGCGTGCGGCTCTTTGCCAACTAGCCTGCGTCACGGGAAATAATATGCTACGGCTTC

541    TGCTTCGTCACCACTTTCACTGCCTGCTTCTGTGCGCGGTTTGGGCAACGCCCTGTCTGG
      |||
532    TGCTTCGTCACCACTTTCACTGCCTGCTTCTGTGCGCGGTTTGGGCAACGCCCTGTCTGG

601    CGTCTCCGTGGTTCACGCTAACGGCGAACCAGAATCCGTCCCCGCCATGGTCTAAACTGA
      |||
592    CGTCTCCGTGGTTCACGCTAACGGCGAACCAGAATCCGTCCCCGCCATGGTCTAAACTGA

661    CGTATCCCAAACCGCATGACGCGGCGACGTTTTACTGTCTTTTCTCTATCCCTCGCCCC
      |||
652    CGTATCCCAAACCGCATGACGCGGCGACGTTTTACTGTCTTTTCTCTATCCCTCGCCCC
```

721 CACGGTCCCCCTCGCAATTCCCGGGGTTCAGCGGGTATCAACGGGTCCCAGTGTGCGA
|||
712 CACGGTCCCCCTCGCAATTCCCGGGGTTCAGCGGGTATCAACGGGTCCCAGTGTGCGA
|||
781 ACGAGACCCTGTATCTGCTGTACAACCGGAAGGCCAGACCTTGGTGGAGAGAAGCTCCA
|||
772 ACGAGACCCTGTATCTGCTGTACAACCGGAAGGCCAGACCTTGGTGGAGAGAAGCTCCA
|||
841 CCTGGGTGAAAAAGGTGATCTGGTATCTGAGCGGTCGCAATCAGACCATCCTCCAACGGA
|||
832 CCTGGGTGAAAAAGGTGATCTGGTATCTGAGCGGTCGCAATCAGACCATCCTCCAACGGA
|||
901 TGCCCCGAACGGCTTCGAAACCGAGCGACGGAACGTGCAGATCAGCGTGGAAGACGCCA
|||
892 TGCCCCGAACGGCTTCGAAACCGAGCGACGGAACGTGCAGATCAGCGTGGAAGACGCCA
|||
961 AGATTTTGGAGCGCACATGGTGCCCAAGCAGACCAAGCTGCTACGTTTCGTGCTCAACG
|||
952 AGATTTTGGAGCGCACATGGTGCCCAAGCAGACCAAGCTGCTACGTTTCGTGCTCAACG
|||
1021 ATGGCACACGTTATCAGATGTGTGTGATGAAACTGGAGAGCTGGGCCCACGTCTTCCGGG
|||
1012 ATGGCACACGTTATCAGATGTGTGTGATGAAACTGGAGAGCTGGGCCCACGTCTTCCGGG
|||
1081 ACTACAGCGTGTCTTTTCAGGTGCGATTGACGTTACCGAGGCCAATAACCAGACTTACA
|||
1072 ACTACAGCGTGTCTTTTCAGGTGCGATTGACGTTACCGAGGCCGATAACCAGACTTACA
|||
1141 CCTTCTGCACCCATCCCAATCTCATCGTTTGAGCCCGTCGCGCGCGCAGGGAATTTTGAA
|||
1132 CCTTCTGCACCCATCCCAATCTCATCGTTTGAGCCCGTCGCGCGCGCAGGGAATTTTGAA
|||
1201 AACCGCGCGTCATGAGTCCCAAAAACCTGACGCCGTTCTTGACGGCGTTGTGGCTGCTAT
|||
1192 AACCGCGCGTCATGAGTCCCAAAAACCTGACGCCGTTCTTGACGGCGTTGTGGCTGCTAT
|||
1261 TGGGTCACAGCCGCGTGCCGCGGGTACGCGCAGAAGAATGTTGCGAATTCATAAACGTCA
|||
1252 TGGGTCACAGCCGCGTGCCGCGGGTACGCGCAGAAGAATGTTGCGAATTCATAAACGTCA
|||
1321 ACCACCCGCCGGAACGCTGTTACGATTTCAAATGTGCAATCGCTTACCGTCGCGTACG
|||
1312 ACCACCCGCCGGAACGCTGTTACGATTTCAAATGTGCAATCGCTTACCGTCGCGTACG
|||
1381 TATTTTCATGATTGTCTGCGTTCTGTGGTGCGTCTGGATTTGTCTCTCGACGTTTCTGAT
|||
1372 TATTTTCATGATTGTCTGCGTTCTGTGGTGCGTCTGGATCTGTCTCTCGACGTTTCTGAT
|||
1441 AGCCATGTTCCATCGACGATCCTCGGAATGCCAGAGTAGATTTTCATGAATCCACAGGC
|||
1432 AGCCATGTTCCATCGACGATCCTCGGAATGCCAGAGTAGATTTTCATGAATCCACAGGC
|||
1501 TGCGGTGTCCGGACGGCGAAGTCTGCTACAGTCCCGAGAAAACGGCTGAGATTCGCGGGA
|||
1492 TGCGGTGTCCGGACGGCGAAGTCTGCTACAGTCCCGAGAAAACGGCTGAGATTCGCGGGA
|||
1561 TCGTCACCACCATGACCCATTCAATTGACACGCCAGGTGCTACACAACAACTGACGAGCT
|||
1552 TCGTCACCACCATGACCCATTCAATTGACACGCCAGGTGCTACACAACAACTGACGAGCT
|||
1621 GCAACTACAATCCGTAAGTCTCTTCCTCGAGGGCCTTACAGCCTATGGGAAAGTAAGACA
|||
1612 GCAACTACAATCC.....

1681 GAGGGACAAAACATCATTAATAAAAAAGTCTAATTTACAGTTTTGTACCCCCCTTCCCC
1625
1741 TCCGTGTTGTAGGTTATACCTCGAAGCTGACGGGCGAATACGCTGCGGCAAAGTGAACGA
|||||
1625GTTATACCTCGAAGCTGACGGGCGAATACGCTGCGGCAAAGTGAACGA
1801 CAAGGCGCAGTACCTGCTGGGCGCCGCTGGCAGCGTTCCCTATCGATGGATCAACCTGGA
|||||
1673 CAAGGCGCAGTACCTGCTGGGCGCCGCTGGCAGCGTTCCCTATCGATGGATCAACCTGGA
|||||
1861 ATACGACAAGATAACCCGGATCGTGGGCGCTGGATCAGTACCTGGAGAGCGTTAAGAAACA
|||||
1733 ATACGACAAGATAGCCCGGATCGTGGGCGCTGGATCAGTACCTGGAGAGCGTTAAGAAACA
|||||
1921 CAAACGGCTGGATGTGTGCCGCGCTAAATGGGCTATATGCTGCAGTGAATAATAAA
|||||
1793 CAAACGGCTGGATGTGTGCCGCGCTAAATGGGCTATATGCTGCAGTGAATAATAAA
|||||

Translation of SEQKlon95-8.txt: HCK-2 (pUL131x1)

1 ATGCGGCTGTGTGCGGTGTGGCTGTCTGTTTGTCTGTGCGCCGTGGTGCTGGGTCAGTGC
1 M R L C R V W L S V C L C A V V L G Q C
61 CAGCGGGAGACCGCAGAAAAAACGATTATTACCGAGTACCGCATTACTGGGACGCGTGC
21 Q R E T A E K N D Y Y R V P H Y W D A C
121 TCTCGCGCGCTGCCTGACCAAACCCGTTACAAGTATGTGGAACAGCTCGTGGACCTCACG
41 S R A L P D Q T R Y K Y V E Q L V D L T
181 TTGAACTACCACTACGATGCGAGCCACGGCTTGGACAACTTTGACGTGCTCAAGAGGTGA
61 L N Y H Y D A S H G L D N F D V L K R *

Figure 11

Comparison RACE clon 95-11 –FIX genomic sequenc

Upper line: SEQFIX UL131-128.txt, from 10 to 1977

Lower line: SEQKlon95-11.txt, from 1 to 1620

SEQFIX UL131-128.txt:SEQKlon95-11.txt identity= 99.57% (1611/1618)
gap=18.24% (361/1979)

```
1      GTCTGCAACATGCGGCTGTGTCTGGGTGTGGCTGTCTGTTTGTCTGTGCGCCGTGGTGCTG
      |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
1      .....ATGCGGCTGTGTCTGGGTGTGGCTGTCTGTTTGTCTGTGCGCCGTGGTGCTG

61     GGTCAAGTCCAGCGGGAGACCGCAG..AAAAAACGATTATTACCGAGTACCGCATTACT
      |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
52     GGTCAAGTCCAGCGGGAGACCGCAGAAAAAACGATTATTACCGAGTACCGCATTACT

119    GGGACGCGTGCTCTCGCGCGCTGCCTGACCAACCCGTTACAAGTATGTGGAACAGCTCG
      |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
112    GGGACGCGTGCTCTCGCGCGCTGCCTGACCAACCCGTTACAAGTATGTGGAACAGCTCG

179    TGGACCTCACGTTGAACTACCACTACGATGCGAGCCACGGCTTGGACAACCTTTGACGTGC
      |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
172    TGGACCTCACGTTGAACTACCACTACGATGCGAGCCACGGCTTGGACAACCTTTGACGTGC

239    TCAAGAGGTGAGGGTACGCGCTAAAGGTGTATGACAACGGGAAGGTAAGGGCGAACGGGT
      |||||||
232    TCAAGAG.....

299    AACGGGTAGGTAACCGCATGGGGTGTGAAATGACGTTTCGGAACCTGTGCTTGCAGAAATCA
      |||||
237    .....AATCA

359    ACGTGACCGAGGTGTCGTTGCTCATCAGCGACTTTAGACGTCAGAACCGTCGCGGCGGCA
      |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
244    ACGTGACCGAGGTGTCGTTGCTCATCAGCGACTTTAGACGTCAGAACCGTCGCGGCGGCA

419    CCAACAAAAGGACCACGTTCAACGCCGCCGGTTCGCTGGCGCCTCACGCCCGGAGCCTCG
      |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
304    CCAACAAAAGGACCACGTTCAACGCCGCCGGTTCGCTGGCGCCTCACGCCCGGAGCCTCG

479    AGTTCAGCGTGCGGCTCTTTGCCAACTAGCCTGCGTCACGGGAAATAATATGCTACGGCT
      |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
364    AGTTCAGCGTGCGGCTCTTTGCCAACTAGCCTGCGTCACGGGAAATAATATGCTACGGCT

539    TCTGCTTCGTCACCACTTTCACTGCCTGCTTCTGTGCGCGGTTTGGGCAACGCCCTGTCT
      |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
424    TCTGCTTCGTCACCACTTTCACTGCCTGCTTCTGTGCGCGGTTTGGGCAACGCCCTGTCT

599    GGCGTCTCCGTGGTTCACGCTAACGGCGAACCAGAATCCGTCCCCGCCATGGTCTAAACT
      |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
484    GGCGTCTCCGTGGTTCACGCTAACGGCGAACCAGAATCCGTCCCCGCCATGGTCTAAACT

659    GACGTATCCCAAACCGCATGACGCGGCGACGTTTTACTGTCTTTTCTCTATCCCTCGCC
      |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
544    GACGTATCCCAAACCGCATGACGCGGCGACGTTTTACTGTCTTTTCTCTATCCCTCGCC
```

1381 CTGCAACTACAATCT.....

Figure 12

Comparison SEQ 128 B - FIX genomic sequence

Upper line: FIX genomic sequence

Lower line: SEQ 128 B

```
5998  ATGAGTCCCAAAAACCTGACGCCGTTCTTGACGGCGTTGTGGCTGCTATTGGGTCACAGC
      |||
1      ATGAGTCCCAAAAACCTGACGCCGTTCTTGACGGCGTTGTGGCTGCTATTGGGTCACAGC

6058  CGCGTGCCGCGGGTACGCGCAGAAGAATGTTGCGAATTCATAAACGTCAACCACCCGCCG
      |||
61     CGCGTGCCGCGGGTACGCGCAGAAGAATGTTGCGAATTCATAAACGTCAACCACCCGCCG

6118  GAACGCTGTTACGATTTCAAATGTGCAATCGCTTCACCGTCGCGTACGTATTTTCATGA
      |||
121    GAACGCTGTTACGATTTCAAATGTGCAATCGCTTCACCGTCGCGTACGTATTTTCATGA

6178  TTGTCTGCGTTCTGTGGTGCGTCTGGATCTGTCTCTCGACGTTTCTGATAGCCATGTTCC
      |||
181    TTGTCTGCGTTCTGTGGTGCGTCTGGATCTGTCTCTCGACGTTTCTGATAGCCATGTTCC

6238  ATCGACGATCCTCGGGAATGCCAGAGTAGATTTTCATGAATCCACAGGCTGCGGTGTCCG
      |||
241    ATCGACGATCCTCGGGAATGCCAGAGTAGATTTTCATGAATCCACAGGCTGCGGTGTCCG

6298  GACGGCGAAGTCTGCTACAGTCCCAGAGAAAACGGCTGAGATTCGCGGGATCGTCACCACC
      |||
301    GACGGCGAAGTCTGCTACAGTCCCAGAGAAAACGGCTGAGATTCGCGGGATCGTCACCACC

6358  ATGACCCATTCATTGACACGCCAGGTGCTACACAACAACTGACGAGCTGCAACTACAAT
      |||
361    ATGACCCATTCATTGACACGCCAGGTGCTACACAACAACTGACGAGCTGCAACTACAAT

6418  CCGTAAGTCTCTTCCTCGAGGGCCTTACAGCCTATGGGAAAGTAAGACAGAGGGACAAAA
      ||
421    CC.....

6478  CATCATTAATAAAAAAGTCTAATTTACAGTTTTGTACCCCCCTTCCCCTCCGTGTTGTA
423    .....

6538  GGTATACCTCGAAGCTGACGGGCGAATACGCTGCGGCAAAGTGAACGACAAGGCGCAGT
      |||
423    .GTTATACCTCGAAGCTGACGGGCGAATACGCTGCGGCAAAGTGAACGACAAGGCGCAGT

6598  ACCTGCTGGGCGCCGCTGGCGGCGTTCCCTATCGATGGATCAACCTGGAATACGACAAGA
      |||
482    ACCTGCTGGGCGCCGCTGGCGGCGTTCCCTATCGATGGATCAACCTGGAATACGACAAGA

6658  TAGCCCGGATCGTGGGCCTGGATCAGTACCTGGAGAGCGTTAAGAAACACAAACGGCTGG
      |||
542    TAGCCCGGATCGTGGGCCTGGATCAGTACCTGGAGAGCGTTAAGAAACACAAACGGCTGG
```



```
6718 ATGTGTGCCGCGCTAAAATGGGCTATATGCTGCAGTGAATAATAAAATGTGTGTTTGTCC
      |||||||||||||||||||||||||||||||||||||||||||||||||||||||
602  ATGTGTGCCGCGCTAAAATGGGCTATATGCTGCAGTGAATAATAAAATGTGTGTTTGTCC

6778 GAAATACGCGTTTTGAGATTTCTG
      ||| |      ||
662  AAAAAAAAAAAAAAAAAAAAAAAAAA
```

Translation of SEQ128 B x 1.txt: HCK-3 (pUL128x1)

```
1      ATGAGTCCCAAAAACCTGACGCCGTTCTTGACGGCGTTGTGGCTGCTATTGGGTCACAGC
1      M S P K N L T P F L T A L W L L L G H S

61     CGCGTGCCGCGGGTACGCGCAGAAGAATGTTGCGAATTCATAAACGTCAACCACCCGCCG
21     R V P R V R A E E C C E F I N V N H P P

121    GAACGCTGTTACGATTTCAAAATGTGCAATCGCTTCACCGTCGCGTACGTATTTTCATGA
41     E R C Y D F K M C N R F T V A Y V F S *
```

Figure 13

Comparison SEQ 128 A - FIX genomic sequence

Upper line: FIX-BAC

Lower line: SEQ128 A

```
5998  ATGAGTCCCAAAAACCTGACGCCGTTCTTGACGGCGTTGTGGCTGCTATTGGGTCACAGC
      ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
1      ATGAGTCCCAAAAACCTGACGCCGTTCTTGACGGCGTTGTGGCTGCTATTGGGTCACAGC

6058  CGCGTGCCGCGGGTACGCGCAGAAGAATGTTGCGAATTCATAAACGTCAACCACCCGCCG
      ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
61     CGCGTGCCGCGGGTACGCGCAGAAGAATGTTGCGAATTCATAAACGTCAACCACCCGCCG

6118  GAACGCTGTTACGATTTCAAATGTGCAATCGCTTCACCGTCGCGTACGTATTTTATGA
      ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
121    GAACGCTGTTACGATTTCAAATGTGCAATCGCTTCACCGTCGC.....

6178  TTGTCTGCGTTCTGTGGTGCCTCTGGATTTGTCTCTCGACGTTTCTGATAGCCATGTTCC

166    .....

6238  ATCGACGATCCTCGGGAATGCCAGAGTAGATTTTCATGAATCCACAGGCTGCGGTGTCCG
      ||||||||||||||
166    .....GCTGCGGTGTCCG

6298  GACGGCGAAGTCTGCTACAGTCCCAGAAAACGGCTGAGATTCGCGGGATCGTCACCACC
      ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
178    GACGGCGAAGTCTGCTACAGTCCCAGAAAACGGCTGAGATTCGCGGGATCGTCACCACC

6358  ATGACCCATTTCATTGACACGCCAGGTCGTACACAACAACTGACGAGCTGCAACTACAAT
      ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
238    ATGACCCATTTCATTGACACGCCAGGTCGTACACAACAACTGACGAGCTGCAACTACAAT

6418  CCGTAAGTCTCTTCCTCGAGGGCCTTACAGCCTATGGGAAAGTAAGACAGAGGGACAAAA
      ||
298    CC.....

6478  CATCATTAATAAAAAAGTCTAATTTACAGTTTTGTACCCCCCTTCCCCTCCGTGTTGTA

300    .....

6538  GGTATACCTCGAAGCTGACGGGCGAATACGCTGCGGCAAAGTGAACGACAAGGCGCAGT
      ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
300    .GTTATACCTCGAAGCTGACGGGCGAATACGCTGCGGCAAAGTGAACGACAAGGCGCAGT

6598  ACCTGCTGGGCGCCGCTGGCAGCGTTCCCTATCGATGGATCAACCTGGAATACGACAAGA
      ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
359    ACCTGCTGGGCGCCGCTGGCAGCGTTCCCTATCGATGGATCAACCTGGAATACGACAAGA

6658  TAACCCGATCGTGGGCCTGGATCAGTACCTGGAGAGCGTTAAGAAACACAAACGGCTGG
      ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
419    TAACCCGATCGTGGGCCTGGATCAGTACCTGGAGAGCGTTAAGAAACACAAACGGCTGG

6718  ATGTGTGCCGCGCTAAAATGGGCTATATGCTGCAGTGAATAATAAAATGTGTGTTTGTCC
      ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
479    ATGTGTGCCGCGCTAAAATGGGCTATATGCTGCAGTGAATAATAAAATGTG
```

Translation of SEQ128 A: HCK-4 (pUL128)

```
1      ATGAGTCCCAAAACCTGACGCCGTTCTTGACGGCGTTGTGGCTGCTATTGGGTCACAGC
1      M S P K [X] L T P F L T A L W L L L G H S

61     CGCGTGCCGCGGGTACGCGCAGAAGAATGTTGCGAATTCATAAACGTCAACCACCCGCCG
21     R V P R V R A E E C C E F I [X] V [X] H P P

121    GAACGCTGTTACGATTTCAAATGTGCAATCGCTTCACCGTCGCACTGCGGTGTCCGGAC
41     E R C Y D F K M C [X] R F T V A L R C P D

181    GGCGAAGTCTGCTACAGTCCCGAGAAACGGCTGAGATTCGCGGGATCGTCACCACCATG
61     G E V C Y S P E K T A E I R G I V T T M

241    ACCCATTCATTGACACGCCAGGTCGTACACAACAACTGACGAGCTGCAACTACAATCTG
81     T H S L T R Q V V H [X] K L T S C [X] Y [X] L

301    TTATACCTCGAAGCTGACGGGCGAATACGCTGCGGCAAAGTGAACGACAAGGCGCAGTAC
101    L Y L E A D G R I R C G K V [X] D K A Q Y

361    CTGCTGGGCGCCGCTGGCAGCGTTCCTATCGATGGATCAACCTGGAATACGACAAGATA
121    L L G A A G S V P Y R W I [X] L E Y D K I

421    ACCCGGATCGTGGGCCTGGATCAGTACCTGGAGAGCGTTAAGAAACACAAACGGCTGGAT
141    T R I V G L D Q Y L E S V K K H K R L D

481    GTGTGCCGCGCTAAAATGGGCTATATGCTGCAGTGA
161    V C R A K M G Y M L Q *
```

Figure 14

Translation of SEQUL130: HCK-5 (pUL130)

```
1      ATGCTACGGCTTCTGCTTCGTCACCACTTTCACTGCCTGCTTCTGTGCGCGGTTTGGGCA
1      M L R L L L R H H F H C L L L C A V W A

61     ACGCCCTGTCTGGCGTCTCCGTGGTTACGCTAACGGCGAACCAGAATCCGTCCCCGCCA
21     T P C L A S P W F T L T A N Q N P S P P

121    TGGTCTAAACTGACGTATCCCAAACCGCATGACGCGGCGACGTTTTACTGTCCTTTTCTC
41     W S K L T Y P K P H D A A T F Y C P F L

181    TATCCCTCGCCCCACGGTCCCCCTCGCAATTCCCGGGGTTCAGCGGGTATCAACGGGT
61     Y P S P P R S P S Q F P G F Q R V S T G

241    CCCGAGTGTGCGAACGAGACCCTGTATCTGCTGTACAACCGGGAAGGCCAGACCTTGGTG
81     P E C R N E T L Y L L Y N R E G Q T L V

301    GAGAGAAGCTCCACCTGGGTGAAAAAGGTGATCTGGTATCTGAGCGGTCGCAATCAGACC
101    E R S S T W V K K V I W Y L S G R N Q T

361    ATCCTCCAACGGATGCCCCGAACGGCTTCGAAACCGAGCGACGGAACGTGCAGATCAGC
121    I L Q R M P R T A S K P S D G N V Q I S

421    GTGGAAGACGCCAAGATTTTTGGAGCGCACATGGTGCCCAAGCAGACCAAGCTGCTACGT
141    V E D A K I F G A H M V P K Q T K L L R

481    TTCGTCGTCAACGATGGCACACGTTATCAGATGTGTGTGATGAAACTGGAGAGCTGGGCC
161    F V V N D G T R Y Q M C V M K L E S W A

541    CACGTCTTCCGGGACTACAGCGTGTCTTTTCAGGTGCGATTGACGTTACCGAGGCCAAT
181    H V F R D Y S V S F Q V R L T F T E A N

601    AACCAGACTTACACCTTCTGCACCCATCCCAATCTCATCGTTTGA
201    N Q T Y T F C T H P N L I V *
```